

California State Water Resources Control Board

Optional Deliverable

Trial format version to be evaluated over three months



Third Party Validation Table Guidelines & Restrictions

Electronic Deliverable Format and Data Dictionary

September 8, 2004

Web site: <http://geotracker.swrcb.ca.gov/>

Table of Contents

ACRONYMS.....	III
1 INTRODUCTION	1
1.1 Document Conventions	2
1.1.1 Table Representation of GeoTracker GEO_VAL File	2
1.1.2 Conventions for Text	3
1.2 Valid Values (VVLs).....	3
2 DATA DELIVERABLE DESCRIPTION	4
3 RELATIONAL FILES FORMAT.....	4
3.1 GEO_VAL: The 3rd Party Validation File	4
3.1.1 File Guidelines and Restrictions:.....	4
3.1.2 The GEO_VAL File Format.....	4
3.1.3 Valid Value Lists for GEO_VAL	5
4.0 SUBMITTAL FILE, RECORD, AND DATA FIELD REQUIREMENTS.....	8
4.1 Submittal File Requirements	8
4.2 Submittal Data Field Requirements	8
4.2.1 Tab Delimited Formats	9
4.2.2 Comma Delimited Format	9
4.3 Procedure for EDD Submittal	9
APPENDIX A: GLOSSARY OF TERMS.....	10

List of Tables

Table 1: GEO_VAL Format.....	4
Table 2: VALQ valid values	5
Table 3: VALRC valid values	5

Acronyms

CSV	Comma Separated Values (also known as comma/quote delimited)
EDD	Electronic Data Deliverable
EDF	Electronic Deliverable Format
GIS	Geographic Information System
GPS	Global Positioning System
LUST	Leaking Underground Storage Tank
PK	Primary Key
RP	Responsible Party
SWRCB	(California) State Water Resources Control Board
UST	Underground Storage Tank
VVL	Valid Value List

1 Introduction

AB2886 (Water Code Sections 13195-13198) requires Responsible Parties to electronically submit compliance data, such as soil or water chemistry analysis, location, and elevation data to the SWRCB Geographical Environmental Information Management System (GeoTracker System).

The GeoTracker GEO_VAL electronic data deliverable (EDD) is an optional reporting table and consists of validation data analytical laboratory data. This deliverable is intended to be produced by the Responsible Party (or their consultant) and to be submitted to the GeoTracker system via Internet. The GeoTracker GEO_VAL data, when uploaded into the SWRCB GeoTracker database, will provide information pertaining to qualifying groundwater and soil testing results.

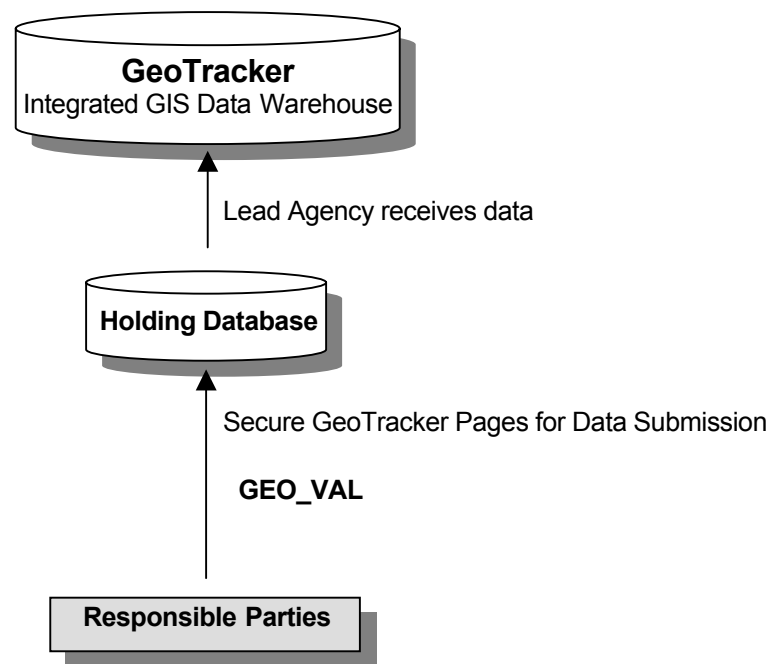


Figure 1: Data Flow for the GeoTracker GEO_VAL Deliverable

1.1 Document Conventions

This document presents the structure and guidelines and restrictions for creating a GeoTracker GEO_VAL electronic data deliverable. The data file for the deliverable is discussed in a level of detail that assists the user in creating an electronic deliverable that meets the criteria of the data standard. Included is a discussion of guidelines and restrictions that apply to files and those that apply to individual fields.

1.1.1 Table Representation of GeoTracker GEO_VAL File

The following table is a representation of the table defining the GeoTracker GEO_VAL

Field Name	Attrb	VVL	RE Q	Dscr. Name	Definition	Guidelines & Restrictions
<i>FIELD1</i>	C18	Yes	Yes	Field 1		Field 1 is a character field with 18 available positions.
<i>FIELD2</i>	D10	No	Yes	Field 2		Field 2 is a date field with an expected format of MM/DD/YYYY.
<i>FIELD3</i>	N5	No	No	Field 3		Field 3 is a numeric field with a total of 5 spaces available for numbers and decimals, with no restriction on the number of digits to the right of the decimal point other than the overall field size.

The “Field Name” is the actual structural name of the field. All primary key fields are shaded within these tables (e.g., *FIELD1* and *FIELD2*). Primary key fields are identified within the tables by shading. **Fields are listed in their structural order within these tables.**

“Attrb” describes the field attributes (type and size). For example:

- C18 is an 18-character field (alphanumeric).
- D10 is a date field with an expected format of MM/DD/YYYY (i.e., 01/01/2001).
- N5 is a numeric field with a total of 5 spaces available for numbers and decimals, with no restriction on the number of digits to the right of the decimal point other than the overall field size (e.g., 12345 or 123.4 or 1.234).

The “VVL” column indicates with a “Yes” or “No” whether the data field requires a valid value code.

The “REQ” column indicates with a “Yes” or “No” whether entry into a field is required.

The “Dscr. Name” column gives the descriptive name of the field.

The “Definition” is a brief definition and/or explanation of the field.

The “Guidelines & Restrictions” describe expectations for entry into the field.

1.1.2 Conventions for Text

Throughout this document, file names are capitalized (e.g., the GEO_XYZ file), and field names are capitalized and italicized (e.g., the *GLOBAL_ID* field). The words “file” and “table” are used interchangeably. The description of each file includes such information as which fields require valid values, which fields require entry for submission, and the file’s primary keys.

1.2 Valid Values (VVLs)

Various data fields in the GeoTracker GEO_VAL electronic deliverable require entry of valid values. Valid values are built-in codes that the format requires for certain fields, such as datum and survey method. The reason for using specific values for these fields is to standardize the data entry, to ensure data consistency, and to help prevent errors. Freely entered data might contain extra spaces, commas, or dashes that would make meaningful data manipulation and thorough or accurate data searches impossible.

Most valid values are abbreviations of common or proper names; hence selecting the correct code is generally straightforward. A list of codes is provided for all fields requiring valid values.

2 Data Deliverable Description

The GeoTracker system relies on a relational database consisting of files related to one another through common (key) fields. These data files are described as relational because the information in one file is related to information in other files, linked through a group of fields called the primary key.

The GEO_VAL deliverable is completed by appending the EDF flat file or the EDF RESULTS table with three additional fields.

3 Relational Files Format

The following Chapter describes the format, data dictionary and guidelines and restrictions associated with the GeoTracker GEO_VAL electronic deliverable.

3.1 GEO_VAL: The 3rd Party Validation File

The GEO_VAL file is an optional file that may be submitted to provide additional qualifying data for analytical results for each groundwater monitoring sample.

3.1.1 File Guidelines and Restrictions:

- Optional fields: *VALPARVAL*, *VALQ*, *VALRC* may be left blank
- Valid Value fields: *VALQ*, *VALRC*, require valid value entries.

3.1.2 The GEO_VAL File Format

These fields are to be appended to the EDF flat or EDF SAMP file.

Table 1: GEO_VAL Format

Field Name	Attrb	VVL	REQ	Dscr. Name	Definition	Guidelines & Restrictions
<i>VALPARVAL</i>	C12	No	Yes	Validation Parameter Value	The adjusted value as deemed appropriate due to validation analysis.	
<i>VALQ</i>	C5	Yes	Yes	Validation Qualifier	The code representing the validation qualifier to be applied to the analytical result.	
<i>VALRC</i>	C5	Yes	Yes	Validation Reason Code	The code representing the reason the qualifier is being applied to the analytical result.	

3.1.3 Valid Value Lists for GEO_VAL

Table 2: VALQ valid values

Code	Description
U	The material was analyzed for, but was not detected above the reporting limit.
J	The analyte was positively identified; the associated numerical value is an estimated quantity (approximate concentration) in the sample.
R	Rejected - The data are unusable due to serious deficiencies in ability to analyze the sample and meet QC criteria (compound may or may not be present). Resampling and reanalysis are necessary for verification.
N	Presumptive evidence of presence of material to make a tentative identification.
UJ	The material was analyzed for, but was not detected above the reporting limit. However, the sample quantitation limit is an estimated quantity and may or may not represent the actual limit of quantitation necessary to accurately and precisely measure the analyte in the sample.
NJ	Presumptive evidence of presence of material to make a tentative identification and the associated numerical value is an estimated quantity or approximate concentration in the sample

Table 3: VALRC valid values

Code	Description
VA	Val. Qual.: Required QA blank not included with samples
VB	Val. Qual.: Analyte present in the blank and the sample
VBX	Val. Qual.: Sample stored at improper temperature
VBV	Val. Qual.: Sample received at improper temperature
VC	Val. Qual.: Calibration nonconformances
VCD	Val. Qual.: Lab Control Sample dup. RPD outside of estab. limits
VCO	Val. Qual.: Analyte conc. detectable, but <5x blank conc.
VCP	Val. Qual.: Analyte conc. detectable, but <10x blank conc.
VCQ	Val. Qual.: Analyte conc. >10x blank conc.
VCR	Val. Qual.: Analyte conc. >5x blank conc.
VDT	Val. Qual.: Diss. metal result > total, beyond std. meth. var.

VDX	Val. Qual.: Value < lowest standard (MQL), but > than MDL
VDZ	Val. Qual.: Positive result is atypical pattern for diesel anal.
VF	Val. Qual.: Compound is common field contaminant
VFB	Val. Qual.: Analyte detected in associated field blank
VFD	Val. Qual.: Field duplicate RPD outside of established limits
VFLB	Val. Qual.: Analyte detected in associated filter blank
VH	Val. Qual.: Holding Time exceedence
VHB	Val. Qual.: Result positively biased
VJ	Val. Qual.: Estimated value
VL	Val. Qual.: Compound is common lab contaminant
VLB	Val. Qual.: Result negatively biased
VLH	Val. Qual.: Lab control sample recoveries above estab. limits
VLL	Val. Qual.: Lab control sample recoveries below estab. limits
VM	Val. Qual.: Nonconformance due to matrix effects
VMB	Val. Qual.: Analyte detected in associated method blank
VMD	Val. Qual.: Matrix Spike dup. RPD outside of established limits
VMH	Val. Qual.: Matrix spike recoveries above established limits
VML	Val. Qual.: Matrix spike recoveries below established limits
VN	Val. Qual.: Identification of compound is tentative
VNB	Val. Qual.: Result bias cannot be determined
VP	Val. Qual.: Sample Chromat. pattern does not match calib. pattern
VPA	Val. Qual.: Possible biogenic interference in detect. hydrocarbs.
VPH	Val. Qual.: Post prep. spike recoveries above established limits
VPL	Val. Qual.: Post prep. spike recoveries below established limits
VQ	Val. Qual.: QA/QC protocols were not met

VQB	Val. Qual.: QA/QC protocols not met for method blank
VQC	Val. Qual.: QA/QC protocols not met for calibration
VQH	Val. Qual.: QA/QC protocols not met for holding times
VQI	Val. Qual.: QA/QC protocols not met for internal standard
VQL	Val. Qual.: QA/QC protocols not met for lab control sample
VQM	Val. Qual.: QA/QC protocols not met for matrix spike/spike dup.
VQN	Val. Qual.: QC data does not exist (hist. data) or is unavailable
VQQ	Val. Qual.: PQL approx. due to QC or matrix effects
VQS	Val. Qual.: QA/QC protocols not met for surrogate recovery
VQT	Val. Qual.: QA/QC protocols not met for instr.12-hr tuning crit.
VQU	Val. Qual.: Non-detect above PQL; final result is now PQL
VR	Val. Qual.: Rejected value
VRB	Val. Qual.: Analyte detected in assoc. equipment rinsate blank
VRL	Val. Qual.: The MQL is above the regulatory limit
VS	Val. Qual.: Sample Receipt nonconformance
VSG	Val. Qual.: A silica gel cleanup procedure was performed
VSH	Val. Qual.: Surrogate recoveries above established limits
VSL	Val. Qual.: Surrogate recoveries below established limits
VSR	Val. Qual.: Semi-quantitative result
VTB	Val. Qual.: Analyte detected in associated trip blank sample
VU	Val. Qual.: Compound analyzed for but not detected
VUB	Val. Qual.: Should be considered non-detect; final result now MDL
VUJ	Val. Qual.: MDL and EQL may be biased low
VYA	Val. Qual.: Independently validated - subject to revision
VYB	Val. Qual.: Specific issue independent validated-subject to revision

VYC	Val. Qual.: Not independent validated-subject to revision
-----	---

4.0 Submittal File, Record, and Data Field Requirements

It is mandatory that the file, record, and data field requirements identified below be adhered to in order to generate acceptable upload file.

4.1 Submittal File Requirements

An EDD may be submitted as a tab-delimited format including: ASCII *.TXT file, Microsoft Excel™ tab delimited *.XLS file, or comma separated value (CSV) delimited ASCII *.TXT file (also known as “comma/quote delimited”).

Each line of data is equivalent to a single record in the data submission. Each record is made up of distinct fields of information. Listed below are the file and record specifications for entering each record of data in its specified file.

- The column heading or field name is not part of the file and should be omitted. Only the data being entered should be in the upload file. Inclusion of column headers (i.e. the field names) in the upload document will cause errors. (Specific instructions for use of Microsoft Excel™ and Microsoft Access™ are detailed below).
- If entering the data via a spreadsheet, such as Microsoft Excel™, enter the first field’s data in the A1 cell. Each cell following A1 should be an additional field of data. When the document is saved as tab-delimited text, excel will save it in the proper format for entry (see instructions below).
- If the record is being created in a text-editor program, there should be no margins and no word-wrapping. (Text-editor programs include Notepad™ and Wordpad™). In each file, every record starts in the farthest left position of “position number 1.”
- Empty rows (records) are not allowed. The first record or row in the file, and every subsequent record or row, must contain valid data.
- Every record within a file must be unique. If, for each key field, a record's data appears exactly the same in another record, these two records are considered to be duplicate records.

4.2 Submittal Data Field Requirements

When producing the fixed or tab delimited formats, data element formats (attributes) must be strictly followed.

- Only authorized codes from the valid value list should be keyed into fields requiring valid values.

- Valid data must always be entered for all required fields.
- Do not add or delete any fields.
- Optional fields where data is not being reported should be left blank (will be converted to unknown).

4.2.1 Tab Delimited Formats

Creating a tab delimited file from a Microsoft Excel™ spreadsheet:

In order to create a file for submittal, create an Excel™ spreadsheet with column headings that match the fields and order required for the GEO_VAL files (specified in this document). Enter valid data for each LUST case in a single row, with each field occupying one cell. Enter the first field's data in the A1 cell. Each cell following A1 should be an additional field of data. Save the spreadsheet as a text document (tab delimited), which is one of the options under file>save as> save as type. At this point, Excel™ will generate a message warning that some formatting may be lost by saving your document in this format. Simply click the "Yes" button, and save the file. By saving as "text (tab delimited)", Excel™ omits the column headers and margins from the document as required for upload. This is the "formatting" that Excel™ warns will be lost when saving the document. This file can be uploaded through the GeoTracker Web-based system.

Creating a tab delimited file from Microsoft Access™

In order to create a file for submittal, first create a database consistent with the GEO_VAL file guidelines. To export data as a tab delimited file, go to file>Export, and set the "Save as type" as "Text Files" and press the "Save All" button. In the new window that will appear, select the "Delimited" option, press "Next", and choose tab delimited. Press "Next" again, and then press "Finish". The new tab-delimited text document will be saved in the directory that was chosen during export.

4.2.2 Comma Delimited Format

CSV delimited format:

For the CSV delimited format, the value entered must be equal to or shorter than the field's length and followed by a comma. For example, when entering a *GLOBAL_ID*, which is a C12 field, if the value to be entered is only C5, in the CSV delimited format it would look like:

"12345","next field entry"

4.3 Procedure for EDD Submittal

Files may be error checked and submitted via the Internet to the GeoTracker system. Dynamic error reports will be generated to inform the user of any data format issues that must be resolved. The files must pass the error-checking program before the data will be submitted to the lead regulatory agency and the GeoTracker database.

Appendix A: Glossary of Terms

Attributes - The format and size attributes of a database field. A field type of C8 is a field that can hold up to eight alphanumeric characters. An N5 field type has a total of 5 spaces available for numbers and decimals, with no restriction on the number of digits to the right of the decimal point other than the overall field size (e.g., 12345 or 123.4 or 1.234). A D10 field type is a date field, and is formatted as MM/DD/YYYY ([month]/[day]/[year]).

Deliverable - A report, data, etc., that is “delivered” to another party, either electronically, or in hard copy format.

EDD (Electronic Data Deliverable) - Information stored in a defined format, accessible via a computer (e.g., stored on diskette, internal hard drive, CD ROM, magnetic tape, etc.).

Field - An area of a table (a column) that contains a particular piece of information. One or more fields make a record. Fields are defined by the attributes of format and size.

File - A named group of electronic data in a defined format.

Guidelines and Restrictions - Information provided to the user regarding data entry, data performance, and data delivery expectations.

Primary Key - A field or set of fields that uniquely identify a record within a table. Key fields within a table define the primary key. Each database record can be uniquely identified using the combination of data fields that make up the primary key.

Record - A line of data (a row) in a table or file made up of distinct fields of information.

Responsible Party- The individual or organization legally responsible for the assessment, monitoring and/or remediation of a contaminated site.

Table - A format for data that allows for data manipulation within a database. Tables are organized with columns and rows of information.

Valid Value - Specially assigned, standardized coded value designating an approved (i.e., “valid”) value for entry into a field in the database.